

CLAIMS

1. (Currently Amended) An [[A]]apparatus for bandwidth managementselectively-queuing data packets for transmission from at least one of a plurality of sources to a given target over a multiplexed communication path, comprising:

a plurality of load shapersload-shaping means responsive to configured to:

maintain a local bandwidth management table comprising a local token count for each of a plurality of classes of source entities;

receive a data packet from a source entity belonging to one of the plurality of classes;

transmit the data packet over a multiplexed communication path if the local token count for the class of the source entity is at least one; and

decrement the local token count for the class of the source entity in the local bandwidth management table in response to the transmission; and

~~data packet communication path transmission requests from one or more managed devices comprising a class, of device(s) and/or application(s), authorized to transmit data up to at least a given minimum bandwidth in a given operational time period; and~~

a BMC (Bandwidth Management Controller) means configured to:

maintain a centralized bandwidth management table comprising a base token count for each of the plurality of classes of source entities, wherein a minimum bandwidth is reserved for each of the plurality of classes of source entities and the base token count increases at a rate corresponding to the minimum bandwidth; and wherein:

the plurality of load shapers is further configured to request a token for the class of the source entity from the Bandwidth Management Controller in response to the transmission; and

the Bandwidth Management Controller is further configured to respond to the request if the base token count for the class of the source entity is at least one by:

providing a token and decrementing the base token count for the class of the source entity.

operable to maintain a table of BW (Bandwidth) authorized for a class and further operable, in response to each request from one of said load shaping means, to supply at least one more token for permitting at least one further data packet to be transmitted at least up to said BW authorized for each given class.

2. (Currently Amended) The apparatus of claim 1, wherein:

the centralized bandwidth management table further comprises a standby token count for each of the plurality of classes of source entities; and

the Bandwidth Management Controller is further configured to respond to the request if the base token count for the class of the source entity is zero and the standby token count for the class of the source entity is at least one by:

providing a token and decrementing the standby token count for the class of the source entity.

Apparatus as claimed in claim 1, comprising, in addition:
communication path usage monitoring means operable to permit said BMC to supply tokens to requesting load shaping means that exceed the class authorized BW, for a given operational period, when the communication path is being utilized at less than a given percentage of capacity.

3. (Currently Amended) A method of ~~bandwidth management~~~~maximizing usage of a~~
~~multiplexed communication path between best efforts data sources and data from at least one of a~~
~~plurality of sources of a given priority class, comprising:~~

maintaining a centralized bandwidth management table comprising a base token count for
each of a plurality of classes of source entities, wherein a minimum bandwidth is reserved for each
of the plurality of classes of source entities and the base token count increases at a rate
corresponding to the minimum bandwidth;

~~maintaining a compilation of information as to classes of data sources authorized to have~~
~~priority in a given time period of operation along with authorized minimum BW (Bandwidth) for~~
~~each class and allowable additional standby BW for each class;~~

maintaining a local bandwidth management table comprising a local token count for each of
the plurality of classes of source entities;

receiving a data packet from a source entity belonging to one of the plurality of classes;

transmitting the data packet over a multiplexed communication path if the local token count
for the class of the source entity is at least one;

decrementing the local token count for the class of the source entity in the local bandwidth
management table in response to the transmission;

requesting a token for the local token count for the class of the source entity in response to
the transmission; and

placing a class ID (identification) request for at least one additional transmission token from
a centralized controller having access to said compilation of information, as each data packet is
placed in a unique class queue, corresponding to said class ID, to be transmitted;

if the base token count for the class of the source entity is at least one, providing the requested token and decrementing the base token count for the class of the source entity,
providing additional transmission tokens, as requested, for each authorized class ID, up to the total of authorized minimum BW and presently allowable standby BW; and
adjusting allowable standby BW for each class as an inverse function of present total communication path utilization for both priority and best efforts data packets.

4 – 13. (Cancelled)

14. (Currently Amended) A method of bandwidth management, communication path activity BW (Bandwidth) management in a multi computer system for use by a plurality of BW managed and BW unmanaged entities wishing to transmit data over said path, comprising:

submitting a request, from a [[BW]]bandwidth managed first entity, for a given bandwidth to an assignment entity;

assigning a unique class identity and a designated allowable [[BW]]bandwidth from said assignment entity;

supplying said assigned unique class identity and designated allowable [[BW]]bandwidth from said assignment entity to a plurality of load shaping entities interconnected to [[said]]a communication path, by said assignment entity, of acceptance of said requested BW by said first entity;

maintaining a centralized count of tokens corresponding to said allowable bandwidth for said class identity;

sending a first data packet[[s]] from said first entity to a first load shaping entity in said plurality of said load shaping entities[[y]] for transmission on the communication path ~~bus~~ attended by said load shaping entity, each of said first data packet[[s]] providing class priority information including said unique identity;

transmitting said first data packet over said communication path by said first load shaping entity if said first load shaping entity has at least one token for said class identity;

removing a token from said first load shaping entity in response to the transmission of said first data packet;

requesting an additional token from the centralized count by the first load shaping entity in response to the transmission of said first data packet; and

allocating a given number N of time slots commensurate with said designated BW, for use over a predetermined number of time slots, for use by said first entity for as long as said first entity continues to supply said data packets for transmission; and

permitting transmission of best efforts data packets over said communication path[[bus]] by unmanaged entities when no managed [[BW]]bandwidth entity data packets await transmission.

15 – 16. (Cancelled)

17. (New) The apparatus of claim 2, wherein the Bandwidth Management Controller is further configured to:

linearly increase the standby token count for each of the plurality of classes of source entities when the communication path is not congested; and

exponentially decrease the standby token count for each of the plurality of source entities when the communication path is congested.

18. (New) The apparatus of claim 1, wherein the local token count for each of the plurality of classes of source entities has a maximum count of two tokens.

19. (New) The apparatus of claim 1, wherein the plurality of load shapers is further configured to maintain a count of outstanding requests for tokens.

20. (New) The method of claim 3, wherein:

the centralized bandwidth management table further comprises a standby token count for each of the plurality of classes of source entities;

and further comprising:

if the base token count for the class of the source entity is zero and the standby token count for the class of the source entity is at least one, providing the requested token and decrementing the standby token count for the class of the source entity.

21. (New) The method of claim 20, further comprising:

linearly increasing the standby token count for each of the plurality of classes of source entities when the communication path is not congested; and

exponentially decreasing the standby token count for each of the plurality of source entities when the communication path is congested.

22. (New) The method of claim 3, wherein the local token count for each of the plurality of classes of source entities has a maximum count of two tokens.

23. (New) The method of claim 3, further comprising maintaining a count of outstanding requests for tokens.

24. (New) A computer program product for bandwidth management, the computer program product embodied on a tangible computer readable medium comprising:

computer code for maintaining a centralized bandwidth management table comprising a base token count for each of a plurality of classes of source entities, wherein a minimum bandwidth is reserved for each of the plurality of classes of source entities and the base token count increases at a rate corresponding to the minimum bandwidth;

computer code for maintaining a local bandwidth management table comprising a local token count for each of the plurality of classes of source entities;

computer code for receiving a data packet from a source entity belonging to one of the plurality of classes;

computer code for transmitting the data packet over a multiplexed communication path if the local token count for the class of the source entity is at least one;

computer code for decrementing the local token count for the class of the source entity in the local bandwidth management table in response to the transmission;

computer code for requesting a token for the local token count for the class of the source entity in response to the transmission; and

computer code for, if the base token count for the class of the source entity is at least one, providing the requested token and decrementing the base token count for the class of the source entity.

25. (New) The computer program product of claim 24, wherein:

the centralized bandwidth management table further comprises a standby token count for each of the plurality of classes of source entities;

and further comprising:

computer code for, if the base token count for the class of the source entity is zero and the standby token count for the class of the source entity is at least one, providing the requested token and decrementing the standby token count for the class of the source entity.

26. (New) The method of claim 25, further comprising:

computer code for linearly increasing the standby token count for each of the plurality of classes of source entities when the communication path is not congested; and

computer code for exponentially decreasing the standby token count for each of the plurality of source entities when the communication path is congested.

27. (New) The method of claim 24, wherein the local token count for each of the plurality of classes of source entities has a maximum count of two tokens.

28. (New) The method of claim 24, further comprising maintaining a count of outstanding requests for tokens.

29. (New) The method of claim 14, comprising, in addition:
normally maintaining at least a predetermined minimum bandwidth for all best efforts data packets queued for transmission.
30. (New) The method of claim 14, wherein:
the communication path is a bus;
the first entity is a computer application being processed simultaneously by:
a first processing unit coupled to the first load shaping entity; and
a second processing unit coupled to a second load shaping entity in said plurality of load shaping entities.
31. (New) The method of claim 30, further comprising:
sending a second data packet from said first entity to the second load shaping entity for transmission on the bus, said data packet providing class priority information including said unique identity;
transmitting said second data packet over said bus by said second load shaping entity if said second load shaping entity has at least one token remaining for said class identity;
removing a token from said second load shaping entity in response to the transmission of said second data packet; and
requesting an additional token from the centralized count by the second load shaping entity in response to the transmission of said second data packet.

32. (New) The method of claim 31, wherein:

said first data packet is transmitted to a first device, said first device being part of the same computer system as said first entity; and

said second data packet is transmitted to a second device, said second device being part of the same computer system as said first entity.